

JOINT DEGREE DISTRIBUTIONS OF PREFERENTIAL ATTACHMENT RANDOM GRAPHS

EROL PEKÖZ, ADRIAN RÖLLIN AND NATHAN ROSS

Boston University, National University of Singapore and University of Melbourne

Abstract

We find a simple representation for the limit distribution of the joint degree counts in proportional attachment random graphs and provide optimal rates of convergence to these limits. The results hold for models with any general initial seed graph and any fixed number of outgoing edges. Optimal rates of convergence to the maximum of the degrees are also derived.

Keywords: Preferential attachment random graph; multicolor Pólya urns, distributional approximation.

1 INTRODUCTION

Preferential attachment random graph models have become extremely popular in the fifteen years since they were studied by Barabási and Albert (1999). In the basic models, nodes are sequentially added to the network over time and attach randomly to existing nodes such that connections to higher degree nodes are more likely. The literature around these models has become too vast to survey but van der Hofstad (2013), Newman (2003) and Newman, Barabási and Watts (2006) provide good overviews.

The most popular models are those similar to Barabási and Albert (1999) in which nodes are added sequentially and attach to exactly one randomly chosen existing node and the chance a new node connects to an existing node is proportional to its degree (see Collovecchio, Cotar and LiCalzi (2013), Krapivsky, Redner and Leyvraz (2000), Rudas, Tóth and Valkó (2007) for results on more general attachment rules). The model is typically generalized to allow for vertices to have $m \geq 1$ initial edges by collapsing m vertices in the one initial edge case into a single vertex (possibly causing loops). The most studied feature of these objects is the distribution of the degrees of the nodes; that is, the proportion of nodes that have degree k as the graph grows large. The basic content of Barabási and Albert (1999) and the rigorous formulation Bollobás, Riordan, Spencer and Tusnády (2001) is that, as $k \rightarrow \infty$, this distribution roughly decays proportional to $k^{-\gamma}$ for some $\gamma > 0$; this is the so-called power law behavior.

In this article we study the joint degree distribution of a fixed collection of vertices for the proportional preferential attachment model where each entering node attaches to exactly $m \geq 1$ nodes. We identify the limiting joint distribution of the scaled degrees for the process started from any initial “seed” graph and provide a rate of convergence to this limit.

To state our results in greater detail we first precisely define the random rules governing the evolution of the models we study. We distinguish between two cases: one that allows for loops and the other that does not; the degree distributions in the two models are different and cannot be reconciled by a simple transformation. We first discuss the case where new nodes initially attach to exactly one existing node.

Let $d_k(n)$ denote the degree of vertex k in the graph $G(n)$. Assume that the seed graph $G(0)$ has s vertices with labels $1, \dots, s$ and with initial degrees d_1, \dots, d_s (note that $d_i = d_i(0)$). We construct $G(n)$ from $G(n-1)$ in two possible ways.

Model N. Given the graph $G(n-1)$ having $s+n-1$ vertices, $G(n)$ is formed by adding vertex $s+n$ and attaching it to vertex k with probability

$$\frac{d_k(n-1)}{\sum_{i=1}^{s+n-1} d_i(n-1)}.$$

Model L. Given the graph $G(n-1)$ having $s+n-1$ vertices, $G(n)$ is formed by adding vertex $s+n$ and attaching it to vertex k with probability

$$\frac{d_k(n-1)}{1 + \sum_{i=1}^{s+n-1} d_i(n-1)},$$

and the probability it forms a loop is equal to $1/(1 + \sum_{i=1}^{s+n-1} d_i(n-1))$.

Note that for Model N if the seed graph is a tree then so is G_n and, in particular, $\sum_{j=1}^{n+s} d_j(n) = 2(s+n-1)$. Similarly, if in Model L the seed graph is a forest with roots having an extra loop then so is G_n for each n and $\sum_{j=1}^{n+s} d_j(n) = 2(s+n)$.

We first state our main result with the simplest seed graph; general seed graphs and more edges are discussed below. Denote by $\text{Ga}(r, \lambda)$ a gamma distribution with shape parameter r and rate parameter λ .

Theorem 1.1. *Assume that either*

- (i) *the seed graph $G(0)$ is a graph consisting of two vertices connected by one edge, and $G(n)$ is formed according to Model N, in which case let $a = 1/2$; or*
- (ii) *the seed graph $G(0)$ is a single vertex with a single edge forming a loop, and $G(n)$ is formed according to Model L, in which case let $a = 1$.*

Fix $r \geq 1$ and denote the scaled degree sequence of the first r vertices of $G(n)$ by

$$W(n) = \frac{1}{2n^{1/2}}(d_1(n), \dots, d_r(n)).$$

Let X_1, \dots, X_r be independent random variables with $X_1 \sim \text{Ga}(a, 1)$ and $X_k \sim \text{Ga}(1, 1)$, $2 \leq k \leq r$, and let

$$Z_k = \sqrt{X_1 + \dots + X_k}, \quad 1 \leq k \leq r,$$

and $Y = (Z_1, Z_2 - Z_1, \dots, Z_r - Z_{r-1})$. Then

$$\sup_K |\mathbb{P}[W(n) \in K] - \mathbb{P}[Y \in K]| \leq \frac{C(r)}{n^{1/2}}$$

for some constant $C(r)$, where the supremum ranges over all convex subsets $K \subset \mathbb{R}^r$.

Remark 1.2. The error rate $n^{-1/2}$ is best-possible since the rate of convergence of a scaled integer valued random variable is bounded from below by the scaling; see Peköz, Röllin and Ross (2013a, Lemma 4.1). Also, a bound on the constant $C(r)$ could in principle (but with much added technicality) be made explicit with our methods; we believe it would grow as a polynomial in r .

Since the sets $K_r(t) = \{\max\{x_1, \dots, x_r\} \leq t\}$ are convex in \mathbb{R}^r , we immediately obtain the following corollary to the theorem.

Corollary 1.3. *With W and Y as in Theorem 1.1 under either (i) or (ii), we have*

$$\sup_{t \geq 0} |\mathbb{P}[\max_{1 \leq k \leq r} W_k(n) \leq t] - \mathbb{P}[\max_{1 \leq k \leq r} Y_k \leq t]| \leq \frac{C(r)}{n^{1/2}}.$$

for some $C(r)$.

Before proceeding to the statement of results for general seed graphs, we make a few remarks. Below we say $X \sim \text{Beta}(a, b)$ if X has density proportional to $x^{a-1}(1-x)^{b-1}$ on $0 < x < 1$.

Remark 1.4. To our knowledge the error bounds of the results above are new and, moreover, the limiting joint degree distributions for the basic preferential attachment graphs of Theorem 1.1 have not been explicitly identified in the literature. However, properties and characterizations of the degrees of fixed vertices are available and can be used to check the result and complete the picture. For example Móri (2005, Theorem 2.1) provides a formula for the joint moments of the degrees of fixed vertices in Model N which may be checked to agree with those of our variables. His work also shows that for $r \geq 1$, the scaled joint degree counts $2(W_1(n), \dots, W_r(n))$ have an almost sure limit $(\zeta_1, \zeta_2, \dots, \zeta_r)$ and so our work above provides a complementary rate of convergence and identifies these limits as

$$\zeta_i = 2(\sqrt{X_1 + \dots + X_i} - \sqrt{X_1 + \dots + X_{i-1}}),$$

where $X_1 \sim \text{Ga}(1/2, 1)$, and X_2, X_3, \dots are i.i.d. distributed as $\text{Ga}(1, 1)$. Moreover, Móri (2005, Lemma 3.4 with $\beta = 0$) shows that the variables

$$\tau_j := \frac{\zeta_1 + \dots + \zeta_{j-1}}{\zeta_1 + \dots + \zeta_j}$$

are $\text{Beta}(2j-1, 1)$ and that $\tau_1, \dots, \tau_r, \zeta_1 + \dots + \zeta_r$ are independent; this result is an easy consequence of our representation and usual beta-gamma algebra facts; see Dufresne (2010). Further, Móri (2005, Theorem 3.1, Lemma 3.3) shows that the variable $\max_{1 \leq k \leq n} W_k(n)$ of Corollary 1.3 converges almost surely to a limit which equals $\max_{i \geq 1} \zeta_i$. With our results this implies that $\max_{1 \leq k \leq n} W_k(n)$ converges to

$$2 \max_{i \geq 1} \left\{ \sqrt{X_1 + \dots + X_i} - \sqrt{X_1 + \dots + X_{i-1}} \right\}.$$

The marginal distribution of the i th vertex in these models can be read from Peköz, Röllin and Ross (2013a, Theorem 1.1, Proposition 2.3). Their Models 1 and 2 are our Models N and L, respectively, and their scaling is a constant times \sqrt{n} . For Model N, they identify the distributional limit of $W_i(n)$ for $i \geq 2$ (note that Vertex 1 and 2 have the same distribution in this model) as $\sqrt{\Gamma_1 B_{1/2, i-3/2}}$, where $\Gamma_1 \sim \text{Ga}(1, 1)$ is independent of $B_{1/2, i-3/2} \sim \text{Beta}(1/2, i-3/2)$. In Model L the analogous limit can be written as $\sqrt{\Gamma_1 B_{1/2, i-1}}$ (interpreting $B_{1/2, 0} = 1$). If $X_{1/2} \sim \text{Ga}(1/2, 1)$ independent of X_1, X_2, \dots i.i.d. distributed as $\text{Ga}(1, 1)$, we have

$$\begin{aligned} \sqrt{X_{1/2} + X_2 + \dots + X_i} - \sqrt{X_{1/2} + X_2 + \dots + X_{i-1}} &\stackrel{\mathcal{D}}{=} \sqrt{\Gamma_1 B_{1/2, i-3/2}}, \\ \sqrt{X_1 + X_2 + \dots + X_i} - \sqrt{X_1 + X_2 + \dots + X_{i-1}} &\stackrel{\mathcal{D}}{=} \sqrt{\Gamma_1 B_{1/2, i-1}}. \end{aligned}$$

These intriguing distributional identities are not easily interpretable, but they can be directly verified by comparing Mellin transforms. It would be of interest to obtain a representation of the joint distributions similar in appearance to the right hand side of these identities.

Remark 1.5. Consider Model L started from a loop. If we write $S_i(n) = \sum_{j=1}^i W_j(n)$ then the theorem and the previous remark imply that for $r \geq 1$, the scaled sums of degree counts

$$(S_1(n), \dots, S_r(n)) \xrightarrow{a.s.} (\sqrt{X_1}, \sqrt{X_1 + X_2}, \dots, \sqrt{X_1 + \dots + X_r}),$$

where the X_i are i.i.d. and have distribution $\text{Ga}(1, 1)$. These are the points of an inhomogeneous Poisson point process with intensity $\frac{1}{2}tdt$ which also arises in Aldous' CRT, Aldous (1991, 1993), and is described around Pitman (2006, Theorem 7.9); see also Peköz, Röllin and Ross (2013c, Remark 2.6). The explicit connection is that if we consider $G(n)$ plus the not yet attached half edge of vertex $n + 1$, then there are $2n + 1$ “degrees” which can be bijectively mapped to a binary tree with $n + 1$ leaves. The bijection is defined through Rémy's algorithm for generating uniformly chosen binary plane trees (see the discussion in Peköz, Röllin and Ross (2013c)). The algorithm begins with a binary tree with two leaves and a root, corresponding to the two starting degrees of the loop and the half edge of the second vertex in Model L. In Rémy's algorithm, leaves are added to the tree by selecting a (possibly internal) vertex uniformly at random and inserting a cherry at this vertex (that is, insert a graph with three vertices and two edges with the “elbow” oriented towards the root of the binary tree), so two vertices (one of which is a leaf) are added at each step and these correspond to the two degrees of each edge added in the preferential attachment model. The number of vertices in the spanning tree of the first k leaves added in Rémy's algorithm is exactly the sum of the degrees of the first k vertices in Model L; for more details see Peköz, Röllin and Ross (2013c, Remark 2.6). In particular, the arguments of Peköz, Röllin and Ross (2013c) show that if we first choose a uniform random binary plane tree with $n \geq 2$ leaves and then $k \leq n$ leaves uniformly at random and fix a labeling $1, \dots, k$, then for $T_j(n)$ defined to be the number of vertices in the spanning tree containing the root and the leaves labeled $1, \dots, j$ we have

$$(S_1(n), \dots, S_k(n)) \stackrel{\mathcal{D}}{=} \frac{1}{2n^{1/2}}(T_1(n), \dots, T_k(n)).$$

Theorem 1.1 provides a rate of convergence of this random vector to its limit which is a multivariate analog of Peköz, Röllin and Ross (2013c, Theorem 2.5(i)).

We can now clearly see the connection to the CRT since according to Aldous (1993), uniform random binary plane trees converge to Brownian CRT, and the number of vertices in the spanning tree of k randomly chosen leaves in a uniform binary tree of n leaves converges to the length of the tree induced by Brownian excursion sampled at k uniform times as per Pitman (2006, Theorem 7.9); see also Pitman (1999).

The following result is a generalization of Theorem 1.1 for general seed graphs. The case where each new node connects to $m > 1$ nodes is discussed afterwards. Write $X \sim \text{GGa}(a, b)$ for a random variable X having the generalized gamma distribution with density proportional to $x^{a-1}e^{-x^b}$.

Theorem 1.6. Fix a seed graph $G(0)$ having s vertices and with initial degree sequence d_1, \dots, d_s , and let $m_k = \sum_{i=1}^k d_i$ for $1 \leq k \leq s$. Assume that either

- (i) $G(n)$ follows Model N, in which case let $a_k = m_s + 2(k - s) + 1$ for $k \geq s$; or
- (ii) $G(n)$ follows Model L, in which case let $a_k = m_s + 2(k - s)$ for $k \geq s$.

Fix $r > s$ and let B_1, \dots, B_{r-1} and Z_r be independent random variables such that

$$B_k \sim \begin{cases} \text{Beta}(m_k, d_{k+1}) & \text{if } 1 \leq k < s, \\ \text{Beta}(a_k, 1) & \text{if } s \leq k \leq r, \end{cases}$$

and $Z_r \sim \text{GGa}(a_r, 2)$. Define the products

$$Z_k = B_k \cdots B_{r-1} Z_r, \quad 1 \leq k < r, \quad Z = (Z_1, \dots, Z_r),$$

and $Y = (Z_1, Z_2 - Z_1, \dots, Z_r - Z_{r-1})$. If we denote the scaled degree sequence of the first r vertices of $G(n)$ by

$$W(n) = \frac{1}{2n^{1/2}}(d_1(n), \dots, d_r(n)),$$

then

$$\sup_K |\mathbb{P}[W(n) \in K] - \mathbb{P}[Y \in K]| \leq \frac{C(r, m_s)}{n^{1/2}}$$

for some constant $C(r, m_s)$, where the supremum ranges over all convex subsets $K \subset \mathbb{R}^r$.

Remark 1.7. We can understand the representations in the limits appearing in Theorems 1.6 and 1.1 through beta-gamma algebra. In the case where the seed graph is a loop, that is, in the setting of Theorem 1.1(ii) and Theorem 1.6(ii) with $s = 1$ and $d_1(0) = 2$, the limit vector Z has the two representations

$$Z_k = \sqrt{X_1 + \cdots + X_k} \quad \text{and} \quad Z_k = B_k \cdots B_{r-1} \sqrt{X_1 + \cdots + X_r}$$

(here using the notation of the theorems and the X_i are independent of the B_i). Using the first relation, the basic beta-gamma algebra implies

$$Z_{r-1} = \sqrt{\frac{X_1 + \cdots + X_{r-1}}{X_1 + \cdots + X_r}} \sqrt{X_1 + \cdots + X_r} \stackrel{\mathcal{D}}{=} \sqrt{V} Z_r,$$

where $V \sim \text{Beta}(r-1, 1)$ is independent of Z_r . A simple (and fortunate) calculation shows that $\sqrt{V} \sim \text{Beta}(2(r-1), 1)$. Continuing in this way, we are able to derive the second representation from the first.

Remark 1.8. The representation of the limit vector in Theorem 1.6 is integral to our approach. It says that in the limit for $k > s$, conditional on coordinate Z_{k+1} , the previous coordinate Z_k is an independent beta variable multiplied by Z_{k+1} . On the other hand, for $k > s$, conditional on the sum of the degrees of the first $k+1$ vertices, say $S_{k+1}(n)$, it is not too difficult to see that the sum of the degrees of the first k vertices, $S_k(n)$ will be distributed as a classical Pólya urn run on the order of $S_{k+1}(n)$ steps (c.f., Lemma 3.1 below). Furthermore, Pólya urns limit to beta variables, $S_k(n) \approx B S_{k+1}(n)$, where B is a beta variable independent of $S_{k+1}(n)$ (c.f., Lemma 3.2 below) and together these conditional distributions specify the joint distribution. The distribution of the degrees of the vertices in the seed graph can be understood in the same way, only now because B_1^2, \dots, B_s^2 are not distributed as beta variables, we lose the representation of Theorem 1.1.

Remark 1.9. From Theorem 1.6 we can easily deduce the following alternative representation for the limiting degrees in the seed graph. Denote by $\text{Dir}(\alpha_1, \dots, \alpha_s)$ the Dirichlet distribution with density proportional to $x_1^{\alpha_1} \cdots x_s^{\alpha_s}$ for $x \in [0, 1]^s$ such that $\sum_{i=1}^s x_i = 1$. Let $X = \text{Dir}(d_1, \dots, d_s)$, and let $Z_s \sim \text{GGa}(m_s + 1, 2)$ under Model N (respectively $Z_s \sim \text{GGa}(m_s, 2)$ under Model L), independent of X . Then Y from Theorem 1.6 satisfies $Y \stackrel{\mathcal{D}}{=} Z_s X$.

Remark 1.10. The seed graph in the theorem above need not be a tree. And also note that a result analogous to Corollary 1.3 holds for the maximum degree started from a general seed graph. Such a result could be useful for proving the main conjecture of Bubeck, Mossel and Rácz (2014), where properties of the maximum degree for star seed graphs are used to show that proportional attachment started with different seed graphs may be separated in total variation even as the number of nodes tends to infinity.

To define a preferential attachment model where each new node attaches initially to $m > 1$ nodes, we follow Bollobás, Riordan, Spencer and Tusnády (2001). The model begins by generating a random graph according to Model L or N with nm nodes, denoted $G(nm)$, and then for each of $i = 1, \dots, n$, collapsing nodes $(i-1)m + 1, \dots, im$ into one node keeping all of the edges (so there may be loops in both models); denote the resulting graph by $G(n)^m$. But with this definition the degree of the i th node in $G(n)^m$ is just the sum of the degrees of nodes $(i-1)m + 1, \dots, im$ in $G(nm)$ and so the limits can be read from Theorems 1.1 and 1.6. Moreover, since a linear transformation of a convex set is convex, the analogous error rates of the theorems in this more general setting also hold. We omit a general formulation due to space and technical considerations, but note that in the setting of Theorem 1.1, summing the limiting distributions of the degrees of adjacent vertices has a particularly simple form due to telescoping. For example for Model L started from a single loop, the joint distributional limits of the scaled degree of nodes $i = 1, \dots, k$ in $G(n)^m$ are given by

$$\sqrt{X_1 + \dots + X_{mi}} - \sqrt{X_1 + \dots + X_{m(i-1)}}$$

where X_1, X_2, \dots are i.i.d. distributed $\text{Ga}(1, 1)$.

The results above rest on relating the degree distributions to an infinite color urn model that generalizes the single color models considered in Janson (2006), Peköz, Röllin and Ross (2013a), Peköz, Röllin and Ross (2013c); urn models frequently appear when studying preferential attachment, see for example Antunović, Mossel and Rácz (2013), Berger, Borgs, Chayes, Saberi et al. (2014), Peköz, Röllin and Ross (2013a), Peköz, Röllin and Ross (2013b), Ross (2013), and Pemantle (2007). In the next section we define the relevant infinite color urn model, state a general approximation result, and then relate it to Theorem 1.6. Section 3 contains the proof of the general approximation result.

2 AN APPROXIMATION THEOREM FOR AN INFINITE-COLOR PÓLYA URN

Consider the following infinite-color urn model. For every positive integer k , let u_k be the number of balls of color k in the urn at time 0 (initial configuration) and assume the urn starts with a finite number of balls so we have $s \geq 1$ such that $u_k \geq 1$ for $k \leq s$ and $u_k = 0$ otherwise. At the n th step, a ball is picked at random from the urn and returned along with an additional ball of the same color, and then a ball having (the new) color $s + n$ is added.

For each $k \geq 1$, let $U_k(n)$ denote the number of balls of color k in the urn after n completed draws; we have $U_k(0) = u_k$. Define the partial sums

$$m_k = u_1 + \dots + u_k, \quad M_k(n) = U_1(n) + \dots + U_k(n).$$

We have the following approximation result for the joint distribution of a finite number of partial sums of color counts.

Theorem 2.1. Fix $r > s$. Let B_1, \dots, B_{r-1} and Z_r be independent random variables such that

$$B_k \sim \begin{cases} \text{Beta}(m_k, u_{k+1}) & \text{if } 1 \leq k < s, \\ \text{Beta}(m_s + 2(k - s) + 1, 1) & \text{if } s \leq k \leq r, \end{cases}$$

and $Z_r \sim \text{GGa}(m_s + 2(r - s) + 1, 2)$. Define

$$Z_k = B_k \cdots B_{r-1} Z_r, \quad 1 \leq k < r, \quad Z = (Z_1, \dots, Z_r),$$

and let

$$W(n) = \frac{1}{2n^{-1/2}} (M_1(n), \dots, M_r(n)).$$

Then there is a constant $C(r, m_s)$, independent of n , such that

$$\sup_K |\mathbb{P}[W(n) \in K] - \mathbb{P}[Z \in K]| \leq \frac{C(r, m_s)}{n^{1/2}} \quad (2.1)$$

for all n , where the supremum ranges over all convex sets $K \subset \mathbb{R}^r$.

Theorem 1.6 easily follows Theorem 2.1 and the next result, which follows from straightforward considerations.

Lemma 2.2. Fix a seed graph $G(0)$ with s vertices and let d_1, \dots, d_s be the initial degree sequence. Consider either the situation of

- (i) Model N for the graph sequence, and an infinite color urn model with initial configuration $(d_1, \dots, d_s, 0, \dots)$, or
- (ii) Model L for the graph sequence, and an infinite color urn model with initial configuration $(d_1, \dots, d_s, 1, 0, \dots)$.

Then, for any r and $n > r$,

$$(d_1(n), \dots, d_r(n)) \stackrel{\mathcal{D}}{=} (U_1(n), \dots, U_r(n)).$$

3 PROOF OF THE URN RESULT

We first need some intermediate results. Denote by $\text{Polya}(b, w; m)$ the distribution of white balls in a classical Pólya urn after m completed draws, starting with b black and w white balls. Matching the notation of Peköz, Röllin and Ross (2013c), denote by $F_{b,w}^{m,1}$ the number of white balls in a generalized Pólya urn after m completed steps starting with b black and w white balls, where one step consists first of a classical Pólya urn step and then the addition of a single black ball.

Lemma 3.1. Let $\ell = k + 1 - s$. If $k \geq s$, we have

$$M_k(n) \sim F_{1, m_s + 2\ell - 1}^{n - \ell, 1}. \quad (3.1)$$

Furthermore, conditionally on $M_{k+1}(n)$, we have

$$M_k(n) \sim \text{Polya}(1, m_s + 2\ell - 1; M_{k+1}(n) - m_s - 2\ell) \quad \text{if } k \geq s, \quad (3.2)$$

and

$$M_k(n) \sim \text{Polya}(u_{k+1}, m_k; M_{k+1}(n) - m_{k+1}) \quad \text{if } 1 \leq k < s. \quad (3.3)$$

Proof. To prove (3.1) note that the number of balls having a color in the set $\{1, \dots, k\}$ is deterministic up to the point where the first ball of color $k+1$ appears in the urn; this is the case after $\ell = k+1-s$ completed draws. At that time we have $M_{k+1}(\ell) = M_k(\ell) + 1 = m_s + 2\ell$. After that, consider all balls of colors $\{1, \dots, k\}$ as ‘white’ balls and all balls of colors $\{k+1, \dots\}$ as ‘black’ balls. The number of ‘white’ balls for the remaining $n - \ell$ steps now behaves exactly like $F_{b,w}^{m,1}$ with $b = 1$, $w = m_s + 2\ell - 1$ and $m = n - \ell$.

To prove (3.2), consider all balls of colors $\{1, \dots, k\}$ as ‘white’ balls and balls of color k as ‘black’ balls. After time ℓ the number of ‘white’ balls now behaves exactly like $\text{Polya}(i, j; m)$ with $i = 1$, $j = M_k(\ell)$, and m being the number of times a ball among colors $\{1, \dots, k+1\}$ was picked after time ℓ , which is just $M_{k+1}(n) - m_s - 2\ell$.

The argument to prove (3.3) is similar and therefore omitted. \square

We will need the following coupling of Pólya urns and beta variables; for related distributional approximation results, see Goldstein and Reinert (2013).

Lemma 3.2 (c.f. Peköz, Röllin and Ross (2013c, Lemma 4.5)). *Fix positive integers i, j and n . There is a coupling (X, Y) with $X \sim \text{Polya}(i, j; n)$ and $Y \sim \text{Beta}(j, i)$, such that*

$$|X - nY| < \frac{i(4j + i + 1)}{2} \quad (3.4)$$

almost surely.

Proof. We use induction over i , and start with the base case $i = 1$; we will prove something slightly more general. Let V_0, \dots, V_{j-1} be independent and uniformly distributed on the interval $[0, 1]$. By a well known representation of the distribution $\text{Beta}(j, 1)$, we can choose

$$Y := \max(V_0, \dots, V_{j-1}).$$

To construct X , first note that

$$\text{Polya}(1, j; m)\{j, \dots, t\} = \prod_{k=0}^{j-1} \frac{t - k}{m + j - k}$$

for all $m \geq 0$ and for $j \leq t \leq j + m$ (see e.g. Feller (1968, Eq. (2.4), p. 121)). For each $m \geq 0$, let

$$N(m) := \max_{0 \leq k \leq j-1} (k + \lceil (m + j - k)V_k \rceil).$$

It is easy to see that the cumulative distribution function of $N(m)$ is that of $\text{Polya}(1, j; m)$ for each m , and that

$$|N(m) - mY| \leq j + 1 \quad \text{for all } m \geq 0. \quad (3.5)$$

Letting $X := N(n)$, (3.4) follows for the case $i = 1$. As a side remark, note that, although $N(m) \sim \text{Polya}(1, j; m)$ for each m , the joint distribution of $(N(0), N(1), \dots)$ is not that of a Pólya urn process!

To prove the inductive step, assume we have constructed $N_{i-1}(0), N_{i-1}(1), \dots$ and Y_{i-1} such that $N_{i-1}(m) \sim \text{Polya}(i-1, j; m)$ for all $m \geq 0$, such that $Y_{i-1} \sim \text{Beta}(j, i-1)$, and such that

$$|N_{i-1}(m) - mY_{i-1}| \leq (i-1)(4j + i)/2 \quad (3.6)$$

for all $m \geq 0$. Now, let $Y'_i \sim \text{Beta}(j, 1)$ be independent of all else and let $N'(0), N'(1), \dots$ be defined and coupled to Y' as in the base case, that is $N'(m) \sim \text{Polya}(1, j; m)$ and $|N'(m) - mY'| \leq j + 1$. Define

$$N_i(m) := N'(N_{i-1}(m) - (j + i - 1)).$$

It is not difficult to see that $N_i(m) \sim \text{Polya}(i, j; m)$. Also, it is not difficult to see that $Y_i := Y_{i-1}Y' \sim \text{Beta}(j, i)$. Noting from (3.5) that for any $y > 0$

$$|N'(m) - yY'| \leq |N'(m) - mY'| + |m - y|Y' \leq (j+1) + |m - y|,$$

we have

$$\begin{aligned} |N_i(m) - mY_i| &= |N'(N_{i-1}(m) - (j+i-1)) - mY_{i-1}Y'| \\ &\leq (j+1) + |N_{i-1}(m) - (j+i-1) - mY_{i-1}| \\ &\leq (j+1) + (j+i-1) + (i-1)(4j+i)/2 \\ &= \frac{i(4j+i+1)}{2}. \end{aligned}$$

This concludes the inductive step, where (3.4) is just the case $m = n$. \square

Lemma 3.3. *For any $k > s$, we have*

$$\mathbb{E}M_k(n) = \frac{(m_s + 2(k-s) + 1)\Gamma((m_s + 2(k-s) + 2)/2)\Gamma(n + \frac{m_s+1}{2})}{\Gamma((m_s + 2(k-s) + 3)/2)\Gamma(n + \frac{m_s}{2})}, \quad (3.7)$$

$$\mathbb{E}\{M_k(n)(M_k(n) + 1)\} = 2n(m_s + 2(k-s) + 1)\left(1 + \frac{m_s + 2(k-s) + 2}{2n}\right), \quad (3.8)$$

and

$$\limsup_{n \rightarrow \infty} n^{1/2} \mathbb{E}M_k(n)^{-1} < \infty \quad (3.9)$$

Proof. The moment expressions follow from the arguments of Peköz, Röllin and Ross (2013c, Lemma 4.1) which says that for $Y \sim F_{1,w}^{t,1}$,

$$\mathbb{E}Y = w \cdot \frac{\Gamma(w/2 + t + 1)\Gamma((w+1)/2)}{\Gamma(w/2 + 1)\Gamma((w+1)/2 + t)}$$

and

$$\mathbb{E}Y(Y+1) = w(w+2t+1).$$

Setting $w = m_s + 2(k-s) + 1$ and $t = n - (k-s+1)$ as per (3.1) yields (3.7) and (3.8).

In order to prove (3.9), let $X \sim F_{1,w+1}^{t,1}$ and $Y \sim F_{1,w}^{t,1}$ where $w = m_s + k - s$. From (3.1) and Peköz, Röllin and Ross (2013c, Lemma 4.2) we have that

$$\mathbb{E}f(X) = \frac{\mathbb{E}\{Yf(Y+1)\}}{\mathbb{E}Y}$$

for any bounded function f ; in particular for the function $f(x) = 1/x$, bounded when $x \geq 1$, we have

$$\mathbb{E}X^{-1} = \mathbb{E}\frac{Y}{(Y+1)\mathbb{E}Y} \leq \frac{1}{\mathbb{E}Y}$$

Using Peköz, Röllin and Ross (2013c, Lemma 4.1) in the same way as for the proof of (3.7), it is clear that $\mathbb{E}Y \asymp n^{1/2}$, from which (3.9) easily follows. \square

Proof of Theorem 2.1. To ease notation, we fix n and drop it in our notation and write, for example, M_k instead of $M_k(n)$. For $k < l$, let

$$K_{k,l} = B_k \cdots B_l, \quad V_{k,l} = K_{k,l-1}W_l, \quad V_{l,l} = W_l.$$

We may assume that B_1, \dots, B_{r-1} and Z_r are all independent of W . Fix a convex subset $K \subset \mathbb{R}^r$, assuming without loss of generality that K is closed, and let h be the indicator function of K . We proceed in two major steps by writing

$$\begin{aligned} & \mathbb{P}[W \in K] - \mathbb{P}[Z \in K] \\ &= \mathbb{E}h(W) - \mathbb{E}h(Z) \\ &= \mathbb{E}\{h(W) - h(V_{1,r}, \dots, V_{r,r})\} + \mathbb{E}\{h(V_{1,r}, \dots, V_{r,r}) - h(Z)\} \\ &=: \mathbb{E}R_1 + \mathbb{E}R_2. \end{aligned}$$

In order to bound R_1 , first write it as the telescoping sum

$$\begin{aligned} R_1 &= h(V_{1,1}, \dots, V_{r,r}) - h(V_{1,r}, V_{2,r}, \dots, V_{r,r}) \\ &= \sum_{k=1}^{r-1} [h(V_{1,k}, \dots, V_{k-1,k}, V_{k,k}, V_{k+1,k+1}, \dots, V_{r,r}) \\ &\quad - h(V_{1,k+1}, \dots, V_{k-1,k+1}, V_{k,k+1}, V_{k+1,k+1}, \dots, V_{r,r})] \\ &=: \sum_{k=1}^{r-1} R_{1,k}. \end{aligned}$$

Fix k , and define a_k and b_k to be the parameters of the Pólya urn in Lemma 3.1, that is,

$$a_k = \begin{cases} 1 & \text{if } k \geq s, \\ u_{k+1} & \text{if } 1 \leq k < s, \end{cases} \quad b_k = \begin{cases} m_s + 2(k - s) + 1 & \text{if } k \geq s, \\ m_k & \text{if } 1 \leq k < s. \end{cases}$$

Conditioning on M_{k+1} , we can use Lemmas 3.1 and 3.2, to conclude that there is a coupling (X_k, Y_k) with

$$X_k \sim \text{Polya}(a_k, b_k; M_{k+1} - c_k), \quad Y_k \sim \text{Beta}(b_k, a_k),$$

such that almost surely,

$$|X_k - (M_{k+1} - c_k)Y_k| \leq \frac{a_k(4b_k + a_k + 1)}{2},$$

where

$$c_k = \begin{cases} m_s + 2(k - 2) + 2 & \text{if } k > s, \\ m_k & \text{if } 1 \leq k < s. \end{cases}$$

Hence, there is a constant $C(k, m_s)$ and a random variable E_k with $|E_k| \leq C(k, m_s)$ almost surely, such that

$$\frac{X_k}{M_{k+1}} = Y_k + \frac{E_k}{M_{k+1}}.$$

Define

$$\begin{aligned} V'_{j,k} &:= \frac{1}{2n^{-1/2}} K_{j,k-1} X_k = K_{j,k-1} \frac{X_k}{M_{k+1}} W_{k+1} \quad \text{for } 1 \leq j < k, \\ V''_{j,k+1} &:= \frac{1}{2n^{-1/2}} K_{j,k-1} Y_k M_{k+1} = K_{j,k-1} Y_k W_{k+1} \quad \text{for } 1 \leq j < k+1, \end{aligned}$$

and

$$\begin{aligned} V'_{j,j} &:= V_{j,j} \quad \text{for } j \geq k, \\ V''_{j,j} &:= V_{j,j} \quad \text{for } j \geq k+1. \end{aligned}$$

Note that $\mathcal{L}(V'_{j,k}) = \mathcal{L}(V_{j,k})$ and $\mathcal{L}(V''_{j,k+1}) = \mathcal{L}(V_{j,k+1})$, hence we can write

$$\begin{aligned} R_{1,k} &= h(V'_{1,k}, \dots, V'_{k-1,k}, V'_{k,k}, V'_{k+1,k+1}, \dots, V'_{r,r}) \\ &\quad - h(V''_{1,k+1}, \dots, V''_{k-1,k+1}, V''_{k,k+1}, V''_{k+1,k+1}, \dots, V''_{r,r}) \\ &= g(Y_k + E_k/M_{k+1}) - g(Y_k) \end{aligned}$$

where

$$g(x) := h(K_{1,k-1}xW_{k+1}, \dots, K_{k-1,k-1}xW_{k+1}, W_{k+1}, \dots, W_r).$$

Let \mathcal{F}_k be the σ -algebra generated by W, B_1, \dots, B_{k-1} and E_k . Since h is the indicator function of a convex set, and since linear transformations of convex sets result again in convex sets, conditional on \mathcal{F}_k , we have that g is of the form $g(x) = \mathbb{I}[a \leq x \leq b]$ for $-\infty \leq a \leq b \leq \infty$. Hence,

$$\begin{aligned} \mathbb{E}(|R_{1,k}| | \mathcal{F}_k) &\leq \mathbb{P}\left[a - \frac{|E_k|}{M_{k+1}} \leq Y_k \leq a + \frac{|E_k|}{M_{k+1}} \mid \mathcal{F}_k\right] + \mathbb{P}\left[b - \frac{|E_k|}{M_{k+1}} \leq Y_k \leq b + \frac{|E_k|}{M_{k+1}} \mid \mathcal{F}_k\right] \\ &\leq 2 \sup_{a \in \mathbb{R}} \mathbb{P}\left[a - \frac{C(k, m_s)}{M_{k+1}} \leq Y_k \leq a + \frac{C(k, m_s)}{M_{k+1}} \mid \mathcal{F}_k\right] \leq \frac{C'(k, m_s)}{M_{k+1}} \end{aligned}$$

for some constant $C'(k, m_s)$. Hence, by Lemma 3.3,

$$\mathbb{E}|R_{1,k}| \leq \mathbb{E} \frac{C'(k, m_s)}{M_{k+1}} \leq \frac{C''(k, m_s)}{n^{1/2}} \quad (3.10)$$

for some $C''(k, m_s)$.

In order to bound R_2 , write

$$\begin{aligned} R_2 &= h((K_{1,r-1}, \dots, K_{r-1,r-1}, 1)W_r) - h((K_{1,r-1}, \dots, K_{r-1,r-1}, 1)Z_r) \\ &= g(W_r) - g(Z_r), \end{aligned}$$

where now

$$g(x) = h((K_{1,r-1}, \dots, K_{r-1,r-1}, 1)x).$$

Given $K_{1,r-1}, \dots, K_{r-1,r-1}$, the function g is again of the same form as in the first part of the proof, so that

$$\mathbb{E}|R_2| \leq 2 d_K(\mathcal{L}(W_r), \mathcal{L}(Z_r)). \quad (3.11)$$

Let $\tilde{\mu}_n = 2n^{1/2}$ and $\mu_n = \sqrt{2\mathbb{E}d_r(n)^2/(m_s + 2(r-s) + 1)}$. According to (3.7) and (3.8), we have $\tilde{\mu}_n/\mu_n = 1 + O(n^{-1/2})$. Since the Kolmogorov distance is scale invariant,

$$\begin{aligned} d_K(\mathcal{L}(W_r), \mathcal{L}(Z_r)) &= d_K(\mathcal{L}(d_r(n)/\mu_n), \mathcal{L}(Z_r \tilde{\mu}_n/\mu_n)) \\ &= d_K(\mathcal{L}(d_r(n)/\mu_n), \mathcal{L}(Z_r)) + d_K(\mathcal{L}(Z_r), \mathcal{L}(Z_r \tilde{\mu}_n/\mu_n)) := R_{2,1} + R_{2,2}. \end{aligned} \quad (3.12)$$

From Peköz, Röllin and Ross (2013c, Theorem 1.2) we have that $R_{2,1} = O(n^{-1/2})$ and, noticing that the density of Z_r is bounded, standard arguments give $R_{2,2} = O(|1 - \tilde{\mu}_n/\mu_n|) = O(n^{-1/2})$. Collecting the bounds (3.10), (3.11) and (3.12), proves (2.1). \square

ACKNOWLEDGMENTS

We thank Jim Pitman for the suggestion to study joint degree distributions for preferential attachment graphs and for pointers to the CRT literature. EP thanks the Department of Mathematics and Statistics at the University of Melbourne for their hospitality during a recent visit when much of this work was completed, supported by a UofM ECR grant. AR was supported by NUS Research Grant R-155-000-124-112.

REFERENCES

- D. J. Aldous (1991). The continuum random tree. I. *Ann. Probab.* **19**, 1–28.
- D. J. Aldous (1993). The continuum random tree. III. *Ann. Probab.* **21**, 248–289.
- T. Antunović, E. Mossel and M. Z. Rácz (2013). Coexistence in preferential attachment networks. *arXiv preprint arXiv:1307.2893*.
- A.-L. Barabási and R. Albert (1999). Emergence of scaling in random networks. *Science* **286**, 509–512.
- N. Berger, C. Borgs, J. T. Chayes, A. Saberi et al. (2014). Asymptotic behavior and distributional limits of preferential attachment graphs. *Ann. Probab.* **42**, 1–40.
- B. Bollobás, O. Riordan, J. Spencer and G. Tusnády (2001). The degree sequence of a scale-free random graph process. *Random Structures Algorithms* **18**, 279–290.
- S. Bubeck, E. Mossel and M. Z. Rácz (2014). On the influence of the seed graph in the preferential attachment model. *arXiv:1401.4849v2*.
- A. Collecchio, C. Cotar and M. LiCalzi (2013). On a preferential attachment and generalized Polya’s urn model. *Ann. Appl. Probab.* **23**, 1219–1253.
- D. Dufresne (2010). G distributions and the beta-gamma algebra. *Electron. J. Probab.* **15**, 2163–2199.
- W. Feller (1968). *An introduction to probability theory and its applications. Vol. I*. John Wiley & Sons Inc., New York, third edition.
- L. Goldstein and G. Reinert (2013). Stein’s method for the Beta distribution and the Pólya-Eggenberger urn. *J. Appl. Probab.* **50**, 1187–1205.
- S. Janson (2006). Limit theorems for triangular urn schemes. *Probab. Theory Related Fields* **134**, 417–452.
- P. L. Krapivsky, S. Redner and F. Leyvraz (2000). Connectivity of growing random networks. *Physical review letters* **85**, 4629.
- T. F. Móri (2005). The maximum degree of the Barabási-Albert random tree. *Combin. Probab. Comput.* **14**, 339–348.
- M. Newman, A.-L. Barabási and D. J. Watts (2006). *The structure and dynamics of networks*. Princeton University Press.
- M. E. J. Newman (2003). The structure and function of complex networks. *SIAM review*.
- E. Peköz, A. Röllin and N. Ross (2013a). Degree asymptotics with rates for preferential attachment random graphs. *Ann. Appl. Probab.* **23**, 1188–1218.
- E. Peköz, A. Röllin and N. Ross (2013b). Total variation error bounds for geometric approximation. *Bernoulli* **19**, 610–632.
- E. Peköz, A. Röllin and N. Ross (2013c). Generalized gamma approximation with rates for urns, walk and trees. *Preprint at* <http://arxiv.org/abs/1309.4183>.
- R. Pemantle (2007). A survey of random processes with reinforcement. *Probab. Surv.* **4**, 1–79. ISSN 1549-5787.
- J. Pitman (1999). Coalescents with multiple collisions. *Ann. Probab.* **27**, 1870–1902.
- J. Pitman (2006). *Combinatorial stochastic processes*, volume 1875 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin. ISBN 978-3-540-30990-1; 3-540-30990-X.
- N. Ross (2013). Power laws in preferential attachment graphs and Stein’s method for the negative binomial distribution. *Adv. Appl. Probab.* **45**, 876–893.
- A. Rudas, B. Tóth and B. Valkó (2007). Random trees and general branching processes. *Random Structures Algorithms* **31**, 186–202.
- R. van der Hofstad (May 2013). Random graphs and complex networks. <http://www.win.tue.nl/~rhofstad/NotesRGCN.pdf>